Automation of host cell protein analysis by mass spectrometry: Increased speed and reproducibility, from sample preparation to report

Carsten P. Sönksen, Ph.D.[1]; Griffin Edward Moran[1]; Jan Bruun Andersen[1]

http://www.novonordisk.co.uk/

Aim

We present a fully automated procedure for MS-based HCP analysis. The following aspects were considered during the development of the procedure:

- Reducing the number of protein cleaning procedures
- Reducing the manual hands-on time
- Increasing the sample capacity
- Increasing the robustness and capacity of the liquid chromatographic method
- Automating the major bottleneck of data analyses and reporting
- Reducing the use of chemicals and environmental impact of the method on employees
- Increasing result quality through system-suitability tests and internal standards

Introduction

Because of the potential risks to patients caused by impurities, detection of host cell proteins (HCPs) present in a drug product is a requirement from regulatory authorities during the development of biopharmaceuticals.

Due to its speed of method development and project flexibility, the ability to assist the downstream development process, and the improved risk assessment enabled by sensitive and accurate HCP identification, mass spectrometry (MS) is increasingly being used for HCP analysis.

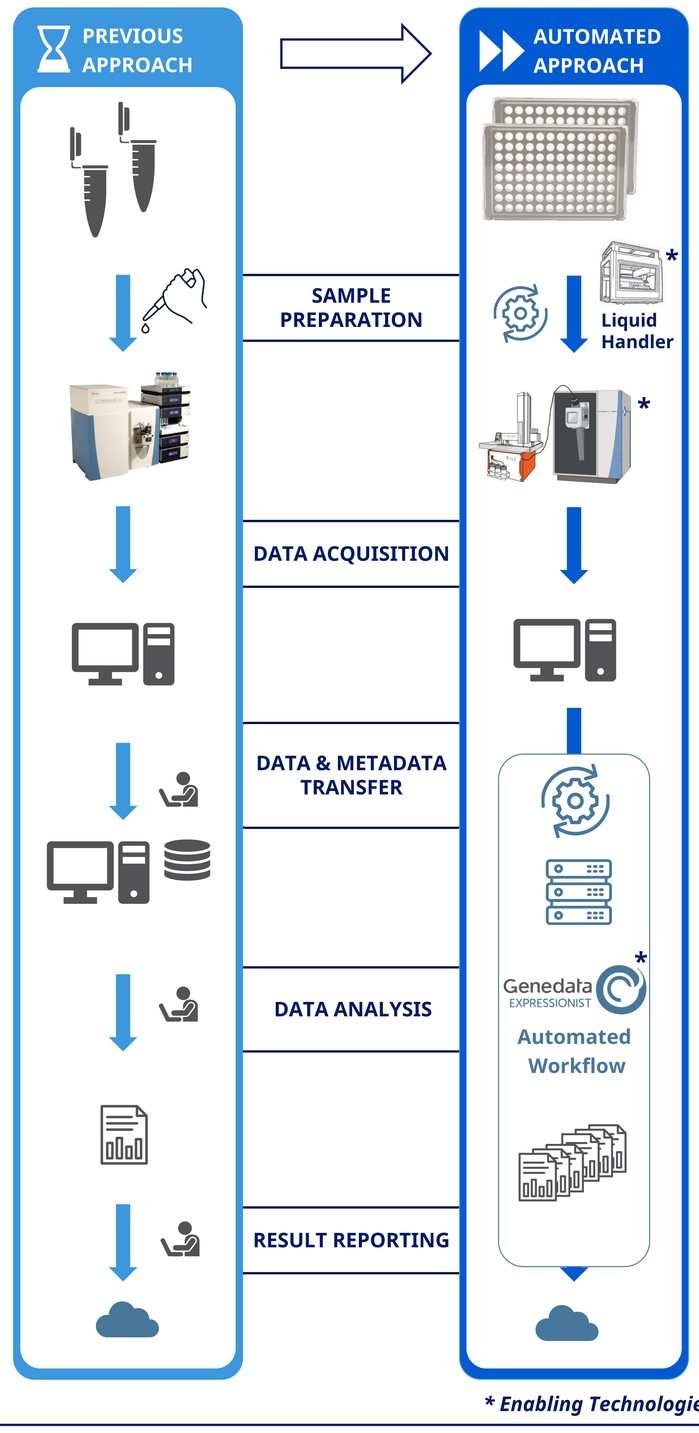
However, MS-based HCP analysis can be a labor-intensive, time-consuming process with limited reproducibility, and may become a bottleneck for complete characterization and rapid production of biopharmaceuticals.

Here we present a fully automated procedure for MS-based HCP analysis.

Methods

- Samples were concentrated by freeze drying.
- HCPs were denatured with urea (8 M), reduced with 10 mM DTT, and alkylated with iodacetamide (100 mM), proteolysis with trypsin (1:50), 4 h @ 37°C.
- LC-MS/MS with Evosep LC (150 µm ID) with 9 or 45 min gradient and Orbitrap mass spectrometer.

Figure 1: Manual and automated HCP analysis approaches

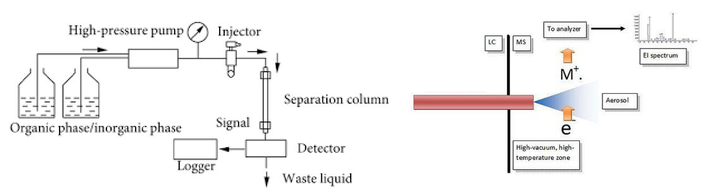


Enabling Technologies

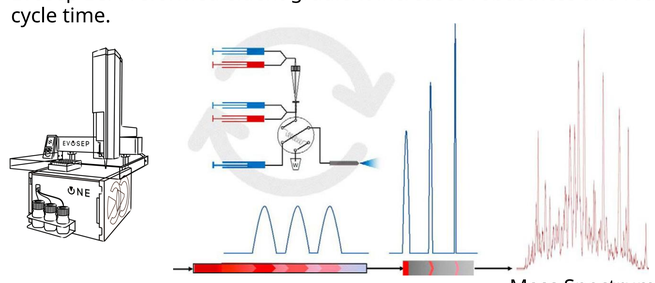
- Liquid handler: Samples prepared in 96-well plates with a Biomek i7 liquid handler using pre-aliquoted reagents.
- Evosep LC: Standardized LC platform designed for robustness and high throughput (see Figure 2).
- Genedata Expressionist: Workflow-based platform for MS data processing, analysis, and reporting (see Figure 3).

Figure 2: Comparison of LC separation methods

- Traditional analytical LC: analytical gradient → long cycle time



- Evosep LC: Preformed eluent gradient increases robustness and reduces cycle time.



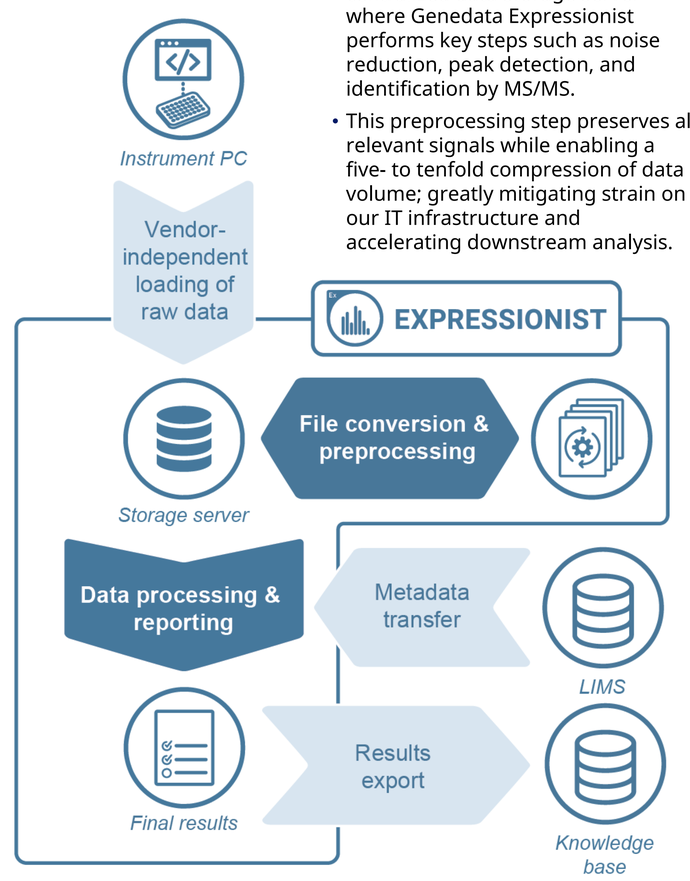
Taken from Reference (1)

Automated Data Processing Workflow

- Genedata Expressionist provides a workflow-based software solution for automated:
  - MS data preprocessing
  - MS and MS/MS data analysis and HCP identification and quantification
  - Customized reporting
  - Integration with external knowledge bases

Figure 3: Genedata Expressionist data processing workflow

- In a fully automated process, the implemented solution transfers experimental raw data to a centralized data storage server where Genedata Expressionist performs key steps such as noise reduction, peak detection, and identification by MS/MS.
- This preprocessing step preserves all relevant signals while enabling a five- to tenfold compression of data volume; greatly mitigating strain on our IT infrastructure and accelerating downstream analysis.

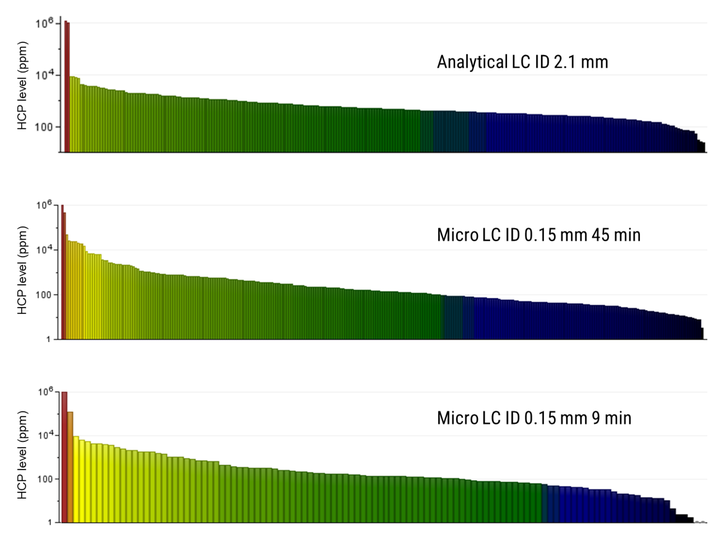


- The seamless integration of Genedata Expressionist into data infrastructure enables information from LIMS and other corporate databases to be leveraged. By automatically retrieving sample metadata—such as sequence identifiers—from a central knowledgebase,
- Genedata Expressionist performs sample-specific data processing, analysis, and reporting. Further sample information—such as name and ID, analysis date, and the desired destination of the final report—is obtained directly from the corporate LIMS identifier.
- MS and MS/MS peaks are automatically detected and submitted to a proteomics search engine (Mascot, Matrix Science) together with the host organism sequence information defined in the sample metadata.

Results

- Successful implementation of automated sample preparation on a Biomek i7 liquid handler increased the sample throughput from 24 to 96 per day.
- Automation resulted in reduction of chemicals applied per sample by a factor of up to thirty and reduced human exposure to toxic chemicals.
- Replacing the labor-intensive analytical LC separation procedure with the Evosep LC resulted in an increased sample throughput from nine samples to 32 or even 160 per day with higher robustness and chromatographic reproducibility.
- The HCP detection dynamic range of 5 orders of magnitude is maintained with the Evosep LC-MS/MS setup (see Figure 4)
- Automation of the data analysis workflow led to reproducible results and the platform could easily cope with the higher sample throughput, thus resolving a previous bottleneck.
- Customization of the reports enabled us to provide a concise yet complete summary of all relevant information on the detected HCPs.

Figure 4: Dynamic range of HCP detection using the given LC conditions



- A customized report is automatically created from the Mascot search results (see Figure 5).
- The report lists all identified contaminants, their abundance (relative to the drug substance) and other information such as molecular weight and pI.
- Links to the Uniprot database enable quick access to information on identified proteins.

Figure 5: Customized HCP report

Name	Abundance [ppm]	UniProt Link	Description	Coverage [%]	MW (kDa)	# Aas	calc. pI
KRT83_HUMAN	246	P78385	Keratin, type II cuticular Hb3	6,97	54,195	493	5,2
KRT85_HUMAN	149	P78386	Keratin, type II cuticular Hb5	4,55	55,802	507	5,5
KRT86_HUMAN	55	O43790	Keratin, type II cuticular Hb6	4,58	53,501	486	6,6
KRT82_HUMAN	83	Q9NSB4	Keratin, type II cuticular Hb2	6,03	56,653	513	8,0
KRT84_HUMAN	37	Q9NSB2	Keratin, type II cuticular Hb4	7,38	64,842	600	5,8

Discussion

- Technologies — such as the Evosep LC — that automate labor-intensive methods provide higher sample throughput while increasing method reproducibility and robustness.
- Implementing end-to-end automation — from sample preparation to data processing, reporting, and integration with our corporate knowledge bases — enabled us to overcome all bottlenecks and create a truly high-throughput HCP analysis platform.
- Environmental aspects — such as reducing human exposure to toxic chemicals and the amount of chemicals required per sample — will be a standard part of analytical automation projects going forward.

Conclusion

By developing and implementing the HCP-MS automation project we achieved the following:

- **Simplified sample purification** — Number of separate procedures reduced from three to one, thereby eliminating the initial manual sorting step.
- **Increased sample throughput** — Automating sample preparation on a liquid handler reduced the hands-on time from four days to 30 minutes, increasing sample capacity from 24 samples to x times 96 samples.
- **Accelerated delivery of high-quality results** — Automated workflow standardized and reduced time required for HCP data processing, analysis, and reporting.
- **Increased method robustness** — Implementing the Evosep LC together with internal standards and system suitability tests increased robustness and reproducibility.
- **Reduced reagent volumes and exposure** — Reagent requirements reduced by a factor of thirty, pre-aliquoting reagents reduces preparation time, reducing employee exposure by a factor of twelve.

[1]Novo Nordisk A/S, Bagsværd, Denmark

The authors acknowledge the assistance of Amy Claydon, Arnd Brandenburg, and Markus Stepath (Genedata).
Presented at SLAS Europe, Thursday May 25, 2023, Brussels, Belgium.

Reference:
(**1**) Bache et al. Mol Cell Proteomics. 2018 Nov;17(11):2284-2296.;